



PCT/FR.2004/000947

REÇU 06 AOUT 2004

OMPI

PCT

# BREVET D'INVENTION

**CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION**

## COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 22 AVR. 2004

Pour le Directeur général de l'Institut  
national de la propriété industrielle  
Le Chef du Département des brevets

Martine PLANCHE

**DOCUMENT DE PRIORITÉ**

PRÉSENTÉ OU TRANSMIS  
CONFORMÉMENT À LA  
RÈGLE 17.1.a) OU b)

INSTITUT  
NATIONAL DE  
LA PROPRIÉTÉ  
INDUSTRIELLE

SIEGE  
26 bis, rue de Saint-Petersbourg  
75800 PARIS cedex 08  
Téléphone : 33 (0)1 53 04 53 04  
Télécopie : 33 (0)1 53 04 45 23  
www.inpi.fr



# BREVET D'INVENTION

## CERTIFICAT D'UTILITE

26bis, rue de Saint-Petersbourg  
75800 Paris Cédex 08  
Téléphone: 01 53.04.53.04 Télécopie: 01.42.94.86.54

Code de la propriété intellectuelle-livre VI

REQUÊTE EN DÉLIVRANCE

DATE DE REMISE DES PIÈCES: 16.04.2003 N° D'ENREGISTREMENT NATIONAL: 0350110 DÉPARTEMENT DE DÉPÔT: 75 DATE DE DÉPÔT: 16.04.2003	BREESE-MAJEROWICZ 3, avenue de l'Opéra 75001 PARIS France
Vos références pour ce dossier: 33458FR	

<b>1 NATURE DE LA DEMANDE</b>			
Demande de brevet			
<b>2 TITRE DE L'INVENTION</b>			
		EMBROUILLAGE, DESEMBROUILLAGE ET DISTRIBUTION SECURISEE DE SEQUENCES AUDIOVISUELLES ISSUES DE CODEURS VIDEOS BASES SUR UN TRAITEMENT PAR ONDELETTES	
<b>3 DECLARATION DE PRIORITE OU REQUETE DU BENEFICE DE LA DATE DE DEPOT D'UNE DEMANDE ANTERIEURE FRANCAISE</b>		Pays ou organisation	Date N°
<b>4-1 DEMANDEUR</b>			
Nom	MEDIALIVE		
Rue	111 avenue Victor Hugo		
Code postal et ville	75116 PARIS		
Pays	France		
Nationalité	France		
Forme juridique	Société anonyme		
N° SIREN	431 742 774		
Code APE-NAF	723Z		
<b>5A MANDATAIRE</b>			
Nom	BREESE-MAJEROWICZ		
Qualité	Org. professionnelle, Pouvoir général		
Rue	3, avenue de l'Opéra		
Code postal et ville	75001 PARIS		
N° de téléphone	01 47 03 67 77		
N° de télécopie	01 47 03 67 78		
Courrier électronique	office@breese.fr		
<b>6 DOCUMENTS ET FICHIERS JOINTS</b>		Fichier électronique	Pages
Texte du brevet		textebrevet.pdf	29
Dessins		dessins.pdf	1
Désignation d'inventeurs		D 22, R 6, AB 1	
Pouvoir général		page 1, figures 1	

<b>7 MODE DE PAIEMENT</b>					
Mode de paiement		Prélèvement du compte courant			
Numéro du compte client		1234			
<b>8 RAPPORT DE RECHERCHE</b>					
Etablissement immédiat					
<b>9 REDEVANCES JOINTES</b>		Devise	Taux	Quantité	Montant à payer
062 Dépôt		EURO	35.00	1.00	35.00
063 Rapport de recherche (R.R.)		EURO	320.00	1.00	320.00
068 Revendication à partir de la 11ème		EURO	15.00	7.00	105.00
Total à acquitter		EURO			460.00

La loi n°78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire.  
Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

**Signé par**

Signataire: FR, FR, Breese-Majerowicz, P. Breese

Emetteur du certificat: DE, DE, D-Trust GmbH, D-Trust for EPO 2.0

**Fonction**

BREESE-MAJEROWICZ (Mandataire 1)

EMBROUILLAGE, DESEMBROUILLAGE ET DISTRIBUTION SECURISEE  
DE SEQUENCES AUDIOVISUELLES ISSUES DE CODEURS VIDEOS  
BASES SUR UN TRAITEMENT PAR ONDELETTES

La présente invention se rapporte au domaine du  
5 traitement de séquences vidéos encodées à l'aide de  
codeurs vidéos basés sur la technologie des ondelettes.

On se propose dans la présente invention de  
fournir un procédé et un système permettant d'embrouiller  
visuellement une séquence vidéo et de recomposer  
10 (désembrouiller) son contenu original à partir d'un flux  
vidéo numérique obtenu par un encodage reposant sur une  
transformée en ondelettes.

La présente invention se rapporte plus  
particulièrement à un dispositif capable de transmettre  
15 de façon sécurisée un ensemble de flux vidéos de haute  
qualité visuelle vers un écran de visualisation type  
écran de télévision et/ou pour être enregistré sur le  
disque dur ou sur tout autre support d'enregistrement  
d'un boîtier reliant le réseau de télétransmission à  
20 l'écran de visualisation tel qu'un écran de télévision ou  
un moniteur d'ordinateur personnel, tout en préservant la  
qualité audiovisuelle mais en évitant toute utilisation  
frauduleuse comme la possibilité de faire des copies  
pirates de films ou de programmes audiovisuels  
25 enregistrés sur le disque dur ou tout autre support  
d'enregistrement du boîtier décodeur. L'invention  
concerne également un système client-serveur et le  
mécanisme de synchronisation entre le serveur qui fournit  
le flux permettant le visionnage du film vidéo numérique  
30 sécurisé et le client qui lit et affiche le flux  
audiovisuel numérique.

Avec les solutions actuelles, il est possible de transmettre des films et des programmes audiovisuels sous forme numérique via des réseaux de diffusion de type hertzien, câble, satellite, etc... ou via des réseaux de télécommunication type DSL (Digital Subscriber Line) ou BLR (boucle locale radio) ou via des réseaux DAB (Digital Audio Broadcasting), etc. Par ailleurs, pour éviter le piratage des œuvres ainsi diffusées, ces dernières sont souvent cryptées ou brouillées par divers moyens bien connus de l'homme de l'art.

Concernant le traitement de séquences vidéo encodées avec la technologie des ondelettes, l'art antérieur connaît le brevet référencé US 6370197 et intitulé « Video Compression Scheme using wavelets » dans lequel les auteurs détaillent une méthode de codage d'une séquence vidéo basée sur une transformée en ondelettes et générant un flux numérique emboîté. Cet art antérieur ne propose aucune méthode de protection du flux et/ou d'embrouillage de la séquence vidéo.

Parmi l'art antérieur concernant la distribution sécurisée de flux audiovisuels multicouches basés sur le principe client - serveur, le brevet US 2001/0053222 A1 propose un procédé et un système pour la protection de flux vidéo encodés avec la norme MPEG-4. Le flux audiovisuel est composé de plusieurs objets audio et vidéo, gérés par une composition scénique. Un des objets du flux vidéo est crypté à l'aide d'une clé générée en quatre étapes de cryptage qui est renouvelée périodiquement. Les objets protégés sont les objets vidéo. L'objet crypté est multiplexé avec les autres objets et le flux entier est envoyé à l'utilisateur. Le flux MPEG-4 est recomposé sur l'équipement destinataire

par le module de décryptage, qui reconstitue le flux vidéo original, à partir du flux vidéo crypté, et en régénérant la clé de cryptage à partir d'informations de cryptage envoyées au préalable et à partir d'informations contenue dans le flux crypté. Etant donné que tout le contenu protégé des objets vidéo se trouve dans le flux envoyé à l'utilisateur, une personne mal intentionnée qui retrouve les clés de cryptage pourra décrypter ce contenu protégé et le visualiser ou diffuser. Cet art antérieur ne résout donc pas entièrement le problème de la sécurisation du flux vidéo.

A l'inverse de la plupart de ces systèmes « classiques » de protection, le procédé conforme à l'invention assure un haut niveau de protection tout en réduisant le volume d'information nécessaire au décodage.

La protection, réalisée de façon conforme à l'invention, est basée sur le principe de suppression et de remplacement de certaines informations codant le signal visuel original par une méthode quelconque, soit : substitution, modification, permutation ou déplacement de l'information. Cette protection est également basée sur la connaissance de la structure du flux binaire à la sortie de l'encodeur visuel basé ondelettes.

5 La présente invention concerne le principe général d'un procédé de sécurisation d'un flux audiovisuel. L'objectif est d'autoriser les services de vidéo à la demande et à la carte à travers tous les réseaux de diffusion et l'enregistrement local dans le boîtier  
10 décodeur numérique de l'utilisateur, ainsi que la visualisation en direct des chaînes de télévision. La solution consiste à extraire et conserver en permanence à l'extérieur de l'habitation de l'utilisateur, en fait dans le

réseau de diffusion et de transmission, une partie du programme audiovisuel enregistré chez le client ou diffusé en direct, cette partie étant primordiale pour visualiser ledit programme audiovisuel sur un écran de télévision ou de type moniteur, mais étant d'un volume très faible par rapport au volume total du programme audiovisuel numérique enregistré chez l'utilisateur ou reçu en temps réel. La partie manquante sera transmise via le réseau de diffusion ou de transmission au moment de la visualisation dudit programme audiovisuel.

Le flux numérique étant séparé en deux parties, la plus grande partie du flux audiovisuel modifié, appelé «flux principal modifié» sera donc transmise via un réseau de diffusion classique alors que la partie manquante appelée «information complémentaire» sera envoyée à la demande via un réseau de télécommunication bande étroite comme les réseaux téléphoniques classiques ou les réseaux cellulaires de type GSM, GPRS ou UMTS ou en utilisant une petite partie d'un réseau de type DSL ou BLR, ou encore en utilisant un sous-ensemble de la bande passante partagée sur un réseau câblé. Le flux audiovisuel est reconstitué sur l'équipement destinataire (décodeur) par un module de synthèse à partir du flux principal modifié et de l'information complémentaire.

L'invention réalise un système de protection, comprenant un module d'analyse - embrouillage et de désembrouillage basé sur un format numérique issu d'un encodage vidéo basé sur des transformées en ondelettes.

Le module d'analyse et d'embrouillage proposé par l'invention repose sur la substitution par des « leurres » ou la modification d'une partie des coefficients issus de la transformation en ondelettes. Le

fait d'avoir enlevé et substitué une partie des données originales du flux vidéo d'origine lors de la génération du flux principal modifié, ne permet pas la restitution dudit flux d'origine à partir des seules données dudit flux principal modifié. De par les caractéristiques de « scalabilité » de la transformée en ondelettes, la notion de « scalabilité » étant définie à partir de l'expression en anglais « scalability », qui caractérise un encodeur capable d'encoder ou un décodeur capable de décoder un ensemble ordonné de flux binaires de façon à produire ou reconstituer une séquence multi couches, plusieurs variantes du procédé d'embrouillage et de désembrouillage sont mises en œuvre et sont illustrées avec des exemples de réalisation.

L'invention concerne selon son acception la plus générale un procédé pour la distribution sécurisée de séquences vidéos selon un format de flux numérique issu d'un encodage basé sur un traitement par ondelettes, constitué de trames (frames) comprenant des blocs contenant des coefficients d'ondelettes décrivant les éléments visuels, caractérisé en ce que l'on procède, avant la transmission à l'équipement client, à une analyse du flux pour générer un flux principal modifié, présentant le format du flux original, et une information complémentaire d'un format quelconque, comportant les informations numériques aptes à permettre la reconstruction desdites trames modifiées, puis à transmettre séparément ledit flux principal modifié et ladite information complémentaire ainsi générés depuis le serveur vers l'équipement destinataire.

De préférence, l'embrouillage est effectué en modifiant des coefficients d'ondelettes appartenant à au



moins une sous-bande temporelle résultant de l'analyse temporelle.

Avantageusement, l'embrouillage est effectué en  
5 modifiant des coefficients d'ondelettes appartenant à moins une sous-bande spatiale résultant de l'analyse spatiale d'une sous bande temporelle.

Selon un mode de réalisation particulier, les  
paramètres pour l'embrouillage dépendent des propriétés  
10 de scalabilité temporelle, de scalabilité spatiale, de scalabilité qualitative, de scalabilité de débit et de scalabilité par régions d'intérêt des flux numériques générés par les codeurs basés ondelettes.

Avantageusement, l'intensité visuelle de la  
15 dégradation des séquences vidéo obtenue est déterminée par la quantité de coefficients d'ondelettes modifiés dans chaque sous-bande spatio-temporelle.

Selon un mode de mise en œuvre particulier, la  
modification des coefficients d'ondelettes est effectué  
20 directement dans le flux binaire.

Selon une variante, la modification des  
coefficients d'ondelettes est effectué avec un décodage  
partiel.

Selon une autre variante, la substitution des  
25 coefficients d'ondelettes est effectuée avec des valeurs aléatoires ou calculées.

De préférence, la durée de l'embrouillage visuel  
obtenu au sein d'une GOF est déterminée en fonction de la  
sous-bande temporelle à laquelle appartiennent les  
30 coefficients d'ondelettes modifiés.

Avantageusement, l'embrouillage visuel obtenu au sein d'une GOF est limité spatialement dans une région d'intérêt de chaque trame.

5 Selon un mode de mise en œuvre particulier, on calcule sur l'équipement destinataire une synthèse d'un flux au format nominal en fonction dudit flux principal modifié et de ladite information complémentaire.

10 Selon un mode de mise en œuvre particulier, la transmission dudit flux principal modifié est réalisée à travers un support matériel distribué physiquement [CD-ROM, DVD, disque dur, carte à mémoire flash].

15 Avantageusement, la transmission de ladite information complémentaire est réalisée à travers un support matériel distribué physiquement [carte à mémoire flash, carte à puce].

20 De préférence, la modification des coefficients d'ondelettes est parfaitement réversible (procédé sans perte) et que le flux numérique reconstitué à partir du flux principal modifié et de l'information complémentaire est strictement identique au flux original.

Selon une variante particulière, la reconstitution d'un flux vidéo désembrouillé est contrôlée en terme de frame-rate, de résolution, de débit et de qualité prédéfinis en fonction des droits de l'utilisateur.

25 Selon une autre variante, la reconstitution du flux vidéo désembrouillé est effectuée de manière progressive par étapes jusqu'à la reconstitution du flux vidéo original.

30 L'invention concerne également un système pour la fabrication d'un flux vidéo pour la mise en œuvre du procédé selon l'une des revendications précédentes, comportant au moins un serveur multimédia contenant les

séquences vidéos originales et caractérisé en ce qu'il comporte un dispositif d'analyse du flux vidéo, un dispositif de séparation du flux vidéo original en un flux principal modifié et en une information  
5 complémentaire en fonction de ladite analyse, et au moins un dispositif sur l'équipement destinataire pour la reconstruction du flux vidéo en fonction dudit flux principal modifié et de ladite information complémentaire.

10

La présente invention sera mieux comprise à la lecture de la description d'un exemple non limitatif de réalisation qui suit, se référant à la figure décrivant l'architecture d'ensemble d'un système pour la mise en  
15 œuvre du procédé selon l'invention.

La protection des flux visuels décrite est élaborée en se basant sur la structure des flux binaires et leurs caractéristiques dues à l'encodage basé ondelettes. Nous rappelons ci-après cette structure.

20

Un codeur vidéo basé sur un traitement par ondelettes réalise une décomposition temporelle et spatiale d'une séquence vidéo initiale afin d'obtenir un ensemble de coefficients d'ondelettes spatio-temporels. Ces coefficients sont ensuite quantifiés puis codés par  
25 un codeur entropique afin de générer un ou plusieurs flux binaires emboîtés qui possèdent des propriétés de scalabilité temporelle et/ou scalabilité spatiale (ou de résolution) et/ou scalabilité qualitative et/ou scalabilité de débit et/ou scalabilité de régions  
30 d'intérêt.

La propriété de scalabilité temporelle est la possibilité de décoder à partir d'un unique ou plusieurs

flux binaires emboîtés, des séquences vidéos dont la fréquence d'affichage des trames (nombre de trames par seconde) est variable.

Dans la littérature bien connue par l'homme de l'art, sont employées les notions en anglais « frames » (trames) et « frames rate » (nombre de trames par seconde), notions qui seront utilisées par la suite pour décrire la présente invention.

La propriété de scalabilité spatiale est la possibilité de décoder à partir d'un unique ou à partir de plusieurs flux binaires emboîtés des séquences vidéos dont la résolution (la taille) spatiale des frames est variable.

La propriété de scalabilité qualitative est la possibilité de décoder à partir d'un unique ou plusieurs flux binaires emboîtés des séquences vidéos dont la qualité visuelle des frames, mesurée selon des critères objectifs et/ou subjectifs, est variable.

La propriété de scalabilité de débit est la possibilité de décoder à partir d'un unique ou plusieurs flux binaires emboîtés des séquences vidéos selon un débit moyen (nombre moyen de bits d'information par seconde) variable.

La propriété de scalabilité sur région d'intérêt est la possibilité de décoder à partir d'un unique ou plusieurs flux binaires emboîtés une ou plusieurs zones ciblées au sein de la séquence vidéo.

A l'encodage, une séquence vidéo originale est segmentée en groupes de N frames successives appelé GOF (Group Of Frames), chaque GOF étant ensuite traité de manière indépendante durant l'encodage. On note un GOF de longueur N  $GOF = (F_0, F_1, \dots, F_{N-1})$ ,  $F_i$  étant les frames pour  $i =$

0,1, 2,... N-1). Les coefficients d'ondelettes spatio-temporels sont générés en deux étapes successives et selon une analyse spatiale et temporelle des frames du GOF.

5           La première étape consiste à effectuer une analyse temporelle des N frames de chaque GOF selon la direction estimée du mouvement (analyse temporelle avec estimation de mouvement), ceci afin de supprimer les redondances temporelles et de concentrer spatialement  
10 dans les frames issues de l'analyse temporelle l'énergie et les informations dues au mouvement. Cette analyse temporelle peut être réalisée selon des résolutions spatiales différentes et après décomposition en ondelettes de chaque frame du GOF, l'estimation de  
15 mouvement étant réalisée indépendamment au sein de chaque sous-bande spatiale (estimation de mouvement multi-résolution).

          La seconde étape consiste à effectuer une analyse spatiale des N frames résultant de l'analyse  
20 temporelle à l'aide d'une décomposition en ondelettes afin de supprimer les redondances spatiales et de concentrer l'énergie due aux discontinuités spatiales présentes dans chaque frame.

          L'analyse temporelle est réalisée en plusieurs  
25 itérations.

          A la première itération, p frames successives du GOF original sont analysées au moyen de filtres ondelettes  $f_L$  et  $f_H$  prédéfinis de longueur p et après estimation et compensation du mouvement pour chaque frame  
30 et par rapport à une ou plusieurs frames dites frames de référence. Généralement  $p=2$  et deux frames successives  $(F_{2i}, F_{2i+1}), i=0, \dots, N/2$ , sont filtrées et engendrent une

frame dite de « basse fréquence » ou « moyenne » et notée  $L_i = f_L(F_{2i}, F_{2i+1})$  et une frame dite de « haute fréquence » ou « différence » et notée  $H_i = f_H(F_{2i}, F_{2i+1})$ . A la première itération de l'étape d'analyse temporelle sont ainsi

5 générés un sous-ensemble  $t-L_1$  de frames de type L de longueur  $N/2$  et un sous-ensemble  $t-H_1$  de frames de type H de longueur  $N/2$ , tels que

$$t-L_1 = (L_0, L_1, \dots, L_{N/2}),$$

$$t-H_1 = (H_0, H_1, \dots, H_{N/2}).$$

10 A chaque itération  $k > 1$  suivante, l'estimation/compensation de mouvement et le filtrage temporel sont itérés sur le sous-ensemble de frames  $t-L_{k-1}$  de type L obtenu à l'itération  $k-1$  et sont générés deux

15 nouveaux sous-ensembles de frames  $t-L_k$  et  $t-H_k$  dont la longueur (nombre de frames) est diminuée d'un facteur 2 par rapport à  $t-L_{k-1}$ . Dans certains cas, l'itération porte aussi sur le sous-ensemble de frames  $t-H_{k-1}$ .

Le nombre total d'itérations durant l'analyse temporelle est noté  $n_T$ . Il est compris entre 1 et  $N/2$ . A

20 l'issue de l'analyse temporelle,  $n_T + 1$  sous-bandes temporelles sont générées.

Par exemple, avec  $N=16$  et  $n_T=4$  :

$$L_i^k = f_L(L_{2i}^{k-1}, L_{2i+1}^{k-1}),$$

$$H_i^k = f_H(L_{2i}^{k-1}, L_{2i+1}^{k-1})$$

25

GOF = (F <sub>0</sub> , F <sub>1</sub> , ..., F <sub>15</sub> )			
t-L <sub>1</sub> = (L <sub>0</sub> , ..., L <sub>7</sub> )			t-H <sub>1</sub> = (H <sub>0</sub> , ..., H <sub>7</sub> )
t-L <sub>2</sub> = (LL <sub>0</sub> , ..., LL <sub>3</sub> )		t-H <sub>2</sub> = (LH <sub>0</sub> , ..., LH <sub>3</sub> )	
t-L <sub>3</sub> = (LLL <sub>0</sub> , LLL <sub>1</sub> )	t-H <sub>3</sub> = (LLH <sub>0</sub> , LLH <sub>1</sub> )		
t-L <sub>4</sub> = (LLLL <sub>0</sub> )	t-H <sub>4</sub> = (LLH <sub>0</sub> )		

L'analyse temporelle d'un GOF de longueur  $N=16$  avec  $n_T=4$  génère ainsi 16 nouvelles frames réparties dans  $n_T+1=4+1=5$  sous-bandes temporelles :

- 5       ▪ Sous-bande  $t-L_4$  : 1 frame de type L :  $LLLL_0$ ,
- Sous-bande  $t-H_4$  : 1 frame de type H :  $LLH_0$ ,
- Sous-bande  $t-H_3$  : 2 frames de type H :  $LLH_0$ ,  
 $LLH_1$ ,
- Sous-bande  $t-H_2$  : 4 frames de type H :  $LH_0, \dots, LH_3$
- 10       ▪ Sous-bande  $t-H_1$  : 8 frames de type H :  $H_0, \dots, H_7$ .

L'analyse spatiale est réalisée ensuite sur chacune des frames appartenant à chaque sous-bande temporelle  $t-L_i$  et  $t-H_i$  : chaque frame est décomposée à l'aide d'une transformée en ondelettes discrète à D  
 15 niveaux, générant ainsi  $3 \times D + 1$  sous-bandes spatiales de coefficients d'ondelettes pour chaque frame. Ces sous-bandes spatiales sont notées  $s-LL_0$ ,  $s-HL_1$ ,  $s-LH_1$ ,  $s-HH_1$ ,  
 $s-HL_2$ ,  $s-LH_2$ ,  $s-HH_2$ , ...,  $s-HL_D$ ,  $s-LH_D$ ,  $s-HH_D$ .

A l'issue des analyses temporelles et  
 20 spatiales,  $(n_T+1) \times (3D+1)$  sous-bandes spatio-temporelles de coefficients d'ondelettes sont disponibles :

- $t-L_{nT}(s-LL_0), t-L_{nT}(s-HL_1), t-L_{nT}(s-LH_1), t-L_{nT}(s-HH_1), \dots, t-L_{nT}(s-HL_D), t-L_{nT}(s-LH_D), t-L_{nT}(s-HH_D),$
- $t-H_{nT}(s-LL_0), t-H_{nT}(s-HL_1), t-H_{nT}(s-LH_1), t-H_{nT}(s-HH_1), \dots, t-H_{nT}(s-HL_D), t-H_{nT}(s-LH_D), t-H_{nT}(s-HH_D),$
- 25       ▪ ...
- $t-H_1(s-LL_0), t-H_1(s-HL_1), t-H_1(s-LH_1), t-H_1(s-HH_1), \dots, t-H_1(s-HL_D), t-H_1(s-LH_D), t-H_1(s-HH_D).$

30       Après une analyse de la structure en sous-bandes décrite précédemment, le module d'analyse et d'embrouillage conforme à l'invention effectuée des

modifications (par permutation et/ou substitution et/ou seuillage) d'un sous-ensemble des coefficients d'ondelettes appartenant à une ou plusieurs sous-bandes spatio-temporelles. Ces modifications introduisent une  
 5 dégradation perceptible visuellement (embrouillage) de la séquence vidéo décodée à partir de ces coefficients modifiés. En fonction du nombre de coefficients modifiés, de leurs localisations dans une sous-bande spatiale, de leurs appartenances aux sous-bandes spatio-temporelles et  
 10 du type de modification, un contrôle de l'étendue spatiale et/ou temporelle de l'embrouillage, ainsi qu'un contrôle de la dégradation due à l'embrouillage sont possibles.

En modifiant des coefficients d'ondelettes dans  
 15 des sous-bandes spatiales d'une sous-bande temporelle  $t-X$  de longueur  $l_x$  (i.e. contenant  $l_x$  frames), les  $N/l_x$  frames consécutives du GOF sont embrouillées.

Le choix du type de sous-bande spatiale à laquelle appartient les coefficients d'ondelettes (s-HL ou s-LH ou  
 20 s-HH) permet de contrôler l'aspect visuel de l'embrouillage : pour la sous-bande s-HL des artefacts de direction verticale apparaissent sur les frames (dégradation des discontinuités spatiales verticales), pour la sous-bande s-LH des artefacts horizontaux  
 25 apparaissent (dégradation des discontinuités spatiales horizontales) et pour la sous-bande s-HH des artefacts de type « damier » apparaissent (dégradations conjointes des discontinuités spatiales horizontales et verticales).

Le choix du niveau de résolution  $r$  auquel  
 30 appartient la sous-bande spatiale (s-LL<sub>r</sub> ou s-HL<sub>r</sub> ou s-LH<sub>r</sub> ou s-HH<sub>r</sub>) permet de contrôler l'étendue spatiale de l'embrouillage engendré par la modification des



coefficients d'ondelettes : plus  $r$  est proche de 0, plus l'étendue spatiale sera importante.

Une modification des coefficients d'ondelettes appartenant à une sous-bande de résolution  $r \geq 0$  génère  
5 un embrouillage qui sera visible sur toutes les frames décodées de résolution spatiales supérieure  $r+1, r+2, \dots, R$ .

La modification des coefficients d'ondelettes spatio-temporels est réalisée après un décodage partiel  
10 du flux binaire généré selon un standard ou une norme ou un algorithme ou un format d'encodage. Une fois la modification effectuée, un ré-encodage des coefficients est réalisé afin de générer un flux binaire de taille  
15 identique et qui respecte la conformité vis à vis du standard ou de la norme ou de l'algorithme ou du format d'encodage ayant généré le flux binaire original.

Avantageusement, des sous-ensembles de bits à l'intérieur même du flux binaire original représentant  
20 les coefficients d'ondelettes spatio-temporels codés sont modifiés sans perturber la conformité du flux vis à vis du standard ou de la norme ou de l'algorithme ou du format d'encodage ayant généré le flux binaire original.

L'embrouillage ainsi généré présente des  
25 propriétés de scalabilité temporelle, spatiale, qualitative, de débit et sur zone d'intérêt.

En fonction du nombre de GOF et/ou du nombre de frames embrouillées au sein d'un GOF, l'embrouillage présente une scalabilité temporelle comprise entre :  
30 « toutes les frames de tous les GOF (embrouillage maximal) » et « aucune frame d'aucun GOF » (séquence non embrouillée).

En fonction des résolutions des sous-bandes spatiales auxquelles appartiennent les coefficients d'ondelettes modifiés, l'embrouillage présente une scalabilité spatiale comprise entre : « toutes les  
5 résolutions sont embrouillées » (i.e. de la résolution  $r=0$  à la résolution  $r=R$ ) et « aucune des résolutions n'est embrouillée ».

En fonction du nombre de coefficients d'ondelettes modifiés et des résolutions des sous-bandes spatiales  
10 auxquelles ils appartiennent, l'embrouillage présente une scalabilité qualitative s'échelonnant de : « la totalité de chaque frame est embrouillée », « certaines régions spatiales de chaque frame sont embrouillées » (régions d'intérêts) et « aucun embrouillage n'est appliqué sur  
15 les frames ».

Réciproquement, le désembrouillage présentera lui-aussi les différentes scalabilités énoncées (temporelle, spatiale, qualitative, de débit et de zone d'intérêt).

On comprendra mieux le principe d'embrouillage et  
20 de désembrouillage se basant sur ces différentes scalabilités à l'aide de l'exemple non limitatif de réalisation préféré qui suit.

Sur le dessin en annexe, la figure représente un mode de réalisation préféré particulier du système  
25 client-serveur conforme à l'invention.

Le flux d'origine (1) est directement sous forme numérique ou sous forme analogique. Dans ce dernier cas, le flux analogique est converti par un codeur basé  
30 ondelettes non représenté en un format numérique (2). Le flux vidéo que l'on souhaite sécuriser (2) est passé à un module d'analyse et d'embrouillage (3) qui va générer un flux principal modifié (5) au format identique au flux

d'entrée (2) en dehors de ce que certains des coefficients ont été remplacés par des valeurs différentes de celles d'origine, et est stocké dans le serveur (6). L'information complémentaire (4), de format  
5 quelconque, est également placée dans le serveur (6) et contient des informations relatives aux éléments des images qui ont été modifiés, remplacés, substitués ou déplacés, et leur valeur ou emplacement dans le flux original.

10 Le flux (5) au format identique au flux original est ensuite transmis, via un réseau haut débit (9) de type hertzien, câble, satellite, etc, au terminal du spectateur (8), et plus précisément sur son disque dur (10). Lorsque le spectateur (8) fait la demande de  
15 visionnage du film présent sur son disque dur (10), deux éventualités sont possibles : soit le spectateur (8) ne possède pas tous les droits nécessaires pour voir le film, dans ce cas, le flux vidéo (5) généré par le module d'embrouillage (3) présent sur le disque dur (10) est  
20 passé au système de synthèse (13), via une mémoire tampon de lecture (11), qui ne le modifie pas et le transmet à l'identique à un lecteur afficheur capable de le décoder (14) et son contenu, dégradé visuellement par le module d'embrouillage (3), est affiché sur l'écran de  
25 visualisation (15). Avantageusement, le flux vidéo (5) généré par le module d'embrouillage (3) est passé directement via un réseau (9) à la mémoire tampon de lecture (11) puis au système de synthèse (13).

30 Soit le serveur décide que le spectateur (8) possède les droits pour visualiser correctement le film. Dans ce cas, le module de synthèse (13) fait une demande de visionnage au serveur (6) contenant l'information

complémentaire nécessaire (4) à la reconstitution de la vidéo originale (2). Le serveur (6) envoie alors via le réseau de télécommunication (7) de type ligne téléphonique analogique ou numérique, DSL (Digital Subscriber Line) ou BLR (Boucle Locale Radio), via des réseaux DAB (Digital Audio Broadcasting) ou via des réseaux de télécommunications mobiles numériques (GSM, GPRS, UMTS), l'information complémentaire (4), permettant la reconstitution de la vidéo originale, de façon à ce que le spectateur (8) puisse la stocker dans une mémoire tampon (12). Le module de synthèse (13) procède alors à la restauration du flux vidéo embrouillé qu'il lit dans sa mémoire tampon de lecture (11), des champs modifiés dont il connaît les positions ainsi que les valeurs d'origine sont restituées, grâce au contenu de l'information complémentaire lue dans la mémoire tampon (12) de désembrouillage. La quantité d'informations contenue dans l'information complémentaire (4) qui est envoyée au module de désembrouillage est spécifique, adaptative et progressive pour chaque spectateur et dépend de ses droits, par exemple utilisation unique ou multiple, droit de faire une ou plusieurs copies privées, retard ou anticipation de paiement. Le niveau (qualité, quantité, type) de l'information complémentaire est déterminé également en fonction de la qualité visuelle requise par l'utilisateur. Le codage vidéo basé ondelettes se caractérisant par les scalabilités précédemment décrites permet la restitution du flux vidéo avec des niveaux de qualité, de résolution, de fréquence de trames différents.

Avantageusement, le flux principal modifié (5) est passé directement via un réseau (9) à la mémoire tampon de lecture (11) puis au module de synthèse (13).

Avantageusement, le flux principal modifié (5) est inscrit (enregistré) sur un support physique comme un disque de type CD-ROM, DVD, disque dur, carte à mémoire flash, etc. (9bis). Le flux principal modifié (5) sera ensuite lu depuis le support physique (9bis) par le lecteur de disque (10bis) du boîtier (8) pour être transmis à la mémoire tampon de lecture (11) puis au module de synthèse (13).

Avantageusement, l'information complémentaire (4) est enregistrée sur un support physique (7bis) de format carte de crédit, constitué par une carte à puce ou une carte à mémoire flash. Cette carte (7bis) sera lue par le module (12) du dispositif (8) qui comprend un lecteur de carte (7ter).

Avantageusement, la carte (7bis) contient les applications et les algorithmes qui seront exécutés par le système de synthèse (13).

Avantageusement, le dispositif (8) est un système autonome, portable et mobile.

Nous allons détailler le fonctionnement du module d'analyse et d'embrouillage (3) illustrant le choix de l'embrouillage effectué. La séquence vidéo originale est segmentée en GOF de  $N=16$  frames. L'analyse temporelle avec  $n_T=4$  itérations génère  $n_T+1=5$  sous-bandes temporelles possédant respectivement :

- Sous-bande  $t-L_4$  : 1 frame de type L :  $LLLL_0$
- Sous-bande  $t-H_4$  : 1 frame de type H :  $LLH_0$
- Sous-bande  $t-H_3$  : 2 frames de type H :  $LLH_0, LLH_1$

▪ Sous-bande  $t-H_2$  : 4 frames de type H :  $LH_0, LH_1, LH_2, LH_3$

▪ Sous-bande  $t-H_1$  : 8 frames de type H :  $H_0, H_1, H_2, H_3, H_4, H_5, H_6, H_7$ .

5        La décomposition en cinq sous bandes temporelles offre la possibilité de reconstituer la séquence vidéo initiale selon cinq différents « frame rate ». Chaque frame de résolution R dans chaque sous-bande temporelle  $t-X$  est ensuite décomposée spatialement par une  
10 transformée en ondelette à  $D=4$  niveaux, ce qui donne la possibilité de reconstituer l'image avec cinq résolutions différentes, générant ainsi pour chacune  $3 \times D + 1 = 13$  sous-bandes spatiales :  $LL_0, LH_1, HL_1, HH_1, LH_2, HL_2, HH_2, LH_3, HL_3, HH_3, LH_4, HL_4, HH_4$ .

15        Suite à un tel encodage, la séquence vidéo peut donc être décodée selon des frame-rates allant de  $1/16 \times fr_0$  à  $fr_0$ ,  $fr_0$  étant le frame-rate de la séquence vidéo originale ainsi que selon  $D+1=5$  résolutions.

20        L'embrouillage de la séquence vidéo est réalisée pour chaque GOF de la manière suivante :

- Dans la sous-bande temporelle  $t-L_4$ , les coefficients d'ondelettes des sous-bandes spatiales  $s-HH_2$  et  $s-HH_3$  résultant de la décomposition spatiale en ondelettes de la frame  $LLL_0$  sont extraits et remplacés  
25 par des valeurs aléatoires ou calculées.

- Dans la sous-bande temporelle  $t-H_3$ , les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  résultant de la décomposition spatiale en ondelettes de la frame  $LLH_0$  sont extraits et remplacés par des valeurs  
30 aléatoires ou calculées.

- Dans la sous-bande temporelle  $t-H_1$ , les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$

résultant de la décomposition en ondelettes de la frame  $LH_0$  sont extraits et remplacés par des valeurs aléatoires ou calculées.

- Dans la sous-bande temporelle  $t-H_0$ , les  
5 coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  résultant de la décomposition en ondelettes de la frame  $H_0$  sont extraits et remplacés par des valeurs aléatoires ou calculées.

La séquence vidéo décodée à partir du flux  
10 principal modifié est donc embrouillée totalement, sauf si elle est décodée à des frame-rates égaux à  $1/16 \times fr_0$  et  $1/8 \times fr_0$  (décodage uniquement à partir respectivement des sous-bandes temporelles  $t-L_4$  et  $(t-L_4, t-H_4)$ ) et à des résolutions spatiales  $r=R/16$  ou  $r=R/8$  (décodage  
15 uniquement à partir respectivement des sous-bandes spatiales  $s-LL_0$ ,  $s-LH_1$ ,  $s-HL_1$  et  $s-HH_1$  qui n'ont pas été modifiées). On effectue ainsi des modifications dans toutes les bandes temporelles, mais pas à toutes les résolutions. Le fait de laisser des résolutions non  
20 modifiées permet de pouvoir reconstituer le flux vidéo à partir du flux embrouillé, mais à une qualité nettement inférieure que celle du flux vidéo original.

Le flux ainsi embrouillé est transmis au client  
(8) à sa demande, le désembrouillage est ensuite réalisé  
25 par exemple en cinq étapes, correspondant à différents niveaux de qualité obtenus après chaque étape de désembrouillage. De cette manière on effectue un désembrouillage d'une ou plusieurs couches de scalabilité, la qualité du film visualisé étant contrôlée  
30 par le serveur en fonction des droits de l'utilisateur et de la qualité requise par lui.

Avantageusement, ce désembrouillage se traduit par une atténuation progressive de la dégradation dans le temps jusqu'à reconstituer le contenu original de haute qualité visuelle. Par exemple, la première étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_2$  pour la sous-bande temporelle  $t-L_4$ . L'embrouillage de la séquence vidéo décodée pour une résolution maximale  $R$  et à un frame-rate maximal de  $fr_0$  est alors moins étendu spatialement et plus concentré autour des discontinuités spatiales de chaque frame. La séquence vidéo décodée pour les frame-rate  $1/16 \times fr_0$  ou  $1/8 \times fr_0$  et pour les résolutions  $r=R/16$ ,  $R/8$ ,  $R/4$  n'est en revanche pas du tout embrouillée. La qualité visuelle du film partiellement désembrouillé est minimale, le film est inexploitable visuellement à sa pleine résolution et à son frame-rate original. Cette étape sert d'identification du serveur d'information complémentaire lors de l'établissement de la connexion.

La deuxième étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s-HH_3$  pour la sous-bande temporelle  $t-L_4$ . L'embrouillage de la séquence vidéo décodée pour une résolution  $R$  et à un frame-rate de  $fr_0$  est alors beaucoup moins étendu spatialement et plus concentré autour des discontinuités spatiales de chaque frame. De plus, la séquence vidéo n'est maintenant embrouillée que sur une durée équivalente à la moitié d'un GOF (8 frames sur 16). La séquence vidéo décodée pour les frame-rate  $1/16 \times fr_0$  ou  $1/8 \times fr_0$  et pour les résolutions  $r=R/16$ ,  $R/8$ ,  $R/4$  n'est pas embrouillée. Cette étape sert à ce que l'utilisateur (8) aperçoive le film vidéo partiellement, afin de



décider s'il veut obtenir des droits pour regarder le film. Après confirmation du client, et en fonction du paiement effectué, sont exécutées les étapes trois, quatre ou cinq. La troisième étape de désembrouillage

5 consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s\text{-}HH_3$  pour la sous-bande temporelle  $t\text{-}H_3$ . La séquence vidéo n'est maintenant embrouillée que sur une durée équivalente au quart d'un GOF (4 frames sur 16). La séquence vidéo décodée pour les frame-rate

10  $1/16 \times fr_0$ ,  $1/8 \times fr_0$ ,  $1/4 \times fr_0$  et pour les résolutions  $r=R/16$ ,  $R/8$ ,  $R/4$  n'est pas embrouillée. Le film peut être regardé visuellement, mais il est de basse qualité.

La quatrième étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande

15 spatiale  $s\text{-}HH_3$  pour la sous-bande temporelle  $t\text{-}H_2$ . La séquence vidéo n'est maintenant embrouillée que sur une durée équivalente au huitième d'un GOF (2 frames sur 16). La séquence vidéo décodée pour les frame-rate  $1/16 \times fr_0$ ,  $1/8 \times fr_0$ ,  $1/4 \times fr_0$ ,  $1/2 \times fr_0$  et pour les résolutions  $r=R/16$ ,

20  $R/8$ ,  $R/4$  n'est pas embrouillée. La qualité visuelle du film restitué est moyenne. Pour ces 3<sup>ième</sup> et 4<sup>ième</sup> étapes de désembrouillage, le film vidéo pleine résolution et pour le frame rate original reste partiellement embrouillé, mais à partir de ces étapes, on peut extraire

25 des flux vidéo avec des résolutions et frame-rate inférieurs à ceux du film original. Cela donne la possibilité de fournir des versions du même flux vidéo de moindre résolution, donc à plus faible prix et de mieux contrôler l'accès.

30 La cinquième étape de désembrouillage consiste à restituer les coefficients d'ondelettes de la sous-bande spatiale  $s\text{-}HH_3$  pour la sous-bande temporelle  $t\text{-}H_1$ . La

séquence vidéo est alors totalement désemprouillée, quelque soit le frame-rate de décodage et la résolution. Le flux vidéo reconstitué est strictement identique au flux vidéo original.

REVENDICATIONS

1 - Procédé pour la distribution sécurisée de  
5 séquences vidéos selon un format de flux numérique issu  
d'un encodage basé sur un traitement par ondelettes,  
constitué de trames (frames) comprenant des blocs  
contenant des coefficients d'ondelettes décrivant les  
éléments visuels, caractérisé en ce que l'on procède,  
10 avant la transmission à l'équipement client, à une  
analyse du flux pour générer un flux principal modifié,  
présentant le format du flux original, et une information  
complémentaire d'un format quelconque, comportant les  
informations numériques aptes à permettre la  
15 reconstruction desdites trames modifiées, puis à  
transmettre séparément ledit flux principal modifié et  
ladite information complémentaire ainsi générés depuis le  
serveur vers l'équipement destinataire.

20 2. Procédé pour la distribution sécurisée de  
séquences vidéos selon la revendication 1, caractérisé en  
ce que l'embrouillage est effectué en modifiant des  
coefficients d'ondelettes appartenant à au moins une  
sous-bande temporelle résultant de l'analyse temporelle.

25

3. Procédé pour la distribution sécurisée de  
séquences vidéos selon l'une des revendications  
précédentes, caractérisé en ce que l'embrouillage est  
effectué en modifiant des coefficients d'ondelettes  
30 appartenant à moins une sous-bande spatiale résultant de  
l'analyse spatiale d'une sous bande temporelle.

4. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que les paramètres pour l'embrouillage dépendent des propriétés de scalabilité temporelle, de scalabilité spatiale, de scalabilité qualitative, de scalabilité de débit et de scalabilité par régions d'intérêt des flux numériques générés par les codeurs basés ondelettes.

5. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que l'intensité visuelle de la dégradation des séquences vidéo obtenue est déterminée par la quantité de coefficients d'ondelettes modifiés dans chaque sous bande spatio-temporelle.

6. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la modification des coefficients d'ondelettes est effectué directement dans le flux binaire.

7. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la modification des coefficients d'ondelettes est effectué avec un décodage partiel.

8. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications précédentes, caractérisé en ce que la substitution des

coefficients d'ondelettes est effectuée avec des valeurs aléatoires ou calculées.

5        9. Procédé pour la distribution sécurisée de  
séquences vidéos selon l'une des revendications  
précédentes, caractérisé en ce que la durée de  
l'embrouillage visuel obtenu au sein d'une GOF est  
déterminée en fonction de la sous-bande temporelle à  
laquelle appartiennent les coefficients d'ondelettes  
10 modifiés.

10. Procédé pour la distribution sécurisée de  
séquences vidéos selon l'une des revendications  
précédentes, caractérisé en ce que l'embrouillage visuel  
15 obtenu au sein d'une GOF est limité spatialement dans une  
région d'intérêt de chaque trame.

11. Procédé pour la distribution sécurisée de  
séquences audiovisuelles selon l'une des revendications  
20 précédentes, caractérisé en ce que l'on calcule sur  
l'équipement destinataire une synthèse d'un flux au  
format nominal en fonction dudit flux principal modifié  
et de ladite information complémentaire.

25        12. Procédé pour la distribution sécurisée de  
séquences audiovisuelles selon l'une des revendications  
précédentes, caractérisé en ce que la transmission dudit  
flux principal modifié est réalisée à travers un support  
matériel distribué physiquement [CD-ROM, DVD, disque dur,  
30 carte à mémoire flash].

13. Procédé pour la distribution sécurisée de séquences audiovisuelles selon l'une des revendications précédentes, caractérisé en ce que la transmission de ladite information complémentaire est réalisée à travers  
5 un support matériel distribué physiquement [carte à mémoire flash, carte à puce].

14. Procédé pour la distribution sécurisée de séquences vidéos selon l'une des revendications  
10 précédentes, caractérisé en ce que la modification des coefficients d'ondelettes est parfaitement réversible (procédé sans perte) et que le flux numérique reconstitué à partir du flux principal modifié et de l'information complémentaire est strictement identique au flux  
15 original.

15. Procédé pour la distribution sécurisée de séquences vidéos selon la revendication 14, caractérisé en ce que la reconstitution d'un flux vidéo désembrouillé  
20 est contrôlée en terme de frame-rate, de résolution, de débit et de qualité prédéfinis en fonction des droits de l'utilisateur.

16. Procédé pour la distribution sécurisée de  
25 séquences vidéos selon la revendication 14 et 15, caractérisé en ce que la reconstitution du flux vidéo désembrouillé est effectuée de manière progressive par étapes jusqu'à la reconstitution du flux vidéo original.

30 17. Système pour la fabrication d'un flux vidéo pour la mise en œuvre du procédé selon l'une des revendications précédentes, comportant au moins un

5 serveur multimédia contenant les séquences vidéos originales et caractérisé en ce qu'il comporte un dispositif d'analyse du flux vidéo, un dispositif de séparation du flux vidéo original en un flux principal modifié et en une information complémentaire en fonction de ladite analyse, et au moins un dispositif sur l'équipement destinataire pour la reconstruction du flux vidéo en fonction dudit flux principal modifié et de ladite information complémentaire.

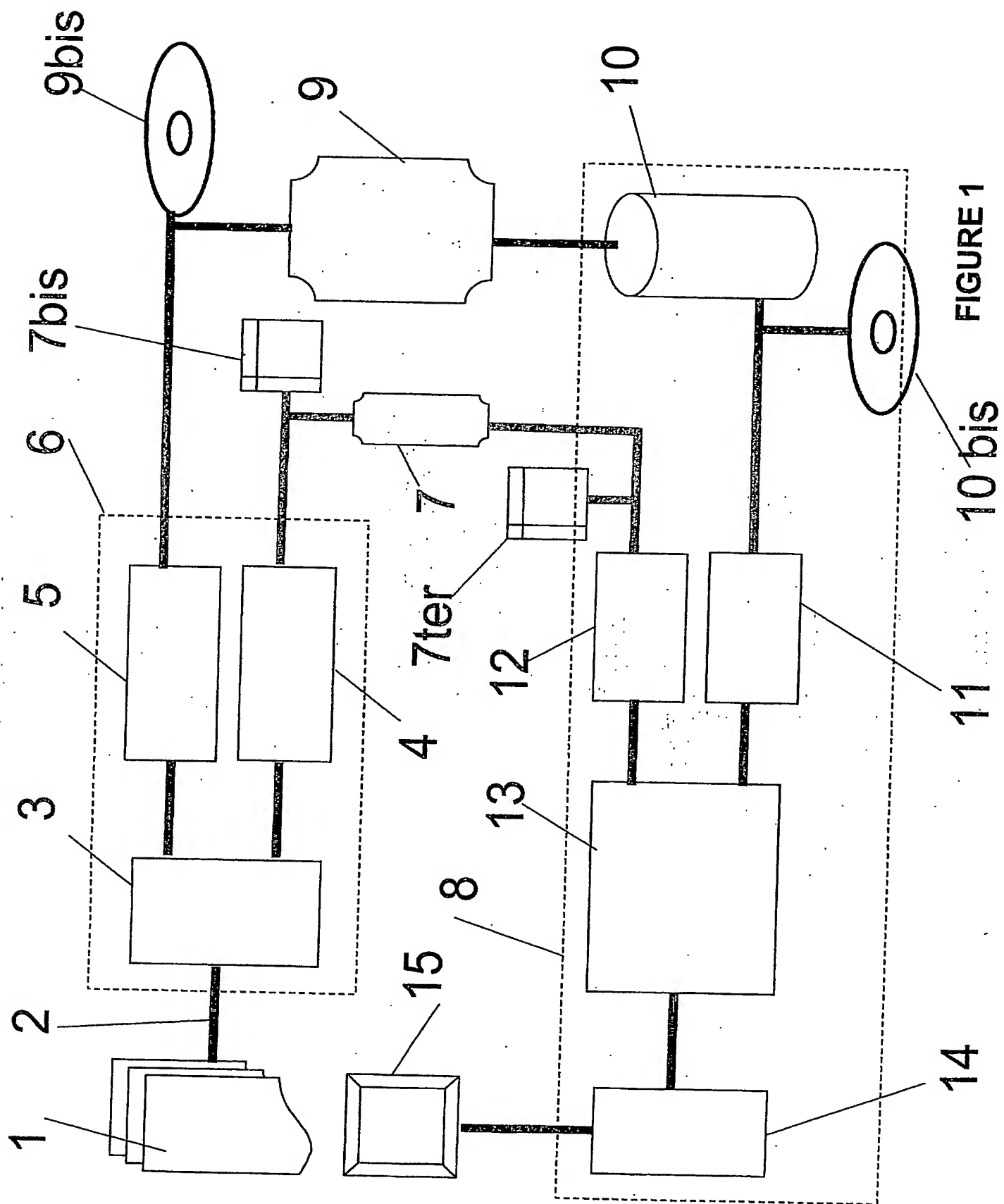


FIGURE 1





## BREVET D'INVENTION CERTIFICAT D'UTILITE

### Désignation de l'inventeur

Vos références pour ce dossier	33458FR
N°D'ENREGISTREMENT NATIONAL	
TITRE DE L'INVENTION	
	EMBROUILLAGE, DESEMBROUILLAGE ET DISTRIBUTION SECURISEE DE SEQUENCES AUDIOVISUELLES ISSUES DE CODEURS VIDEOS BASES SUR UN TRAITEMENT PAR ONDELETTES
LE(S) DEMANDEUR(S) OU LE(S) MANDATAIRE(S):	
DESIGNE(NT) EN TANT QU'INVENTEUR(S):	
Inventeur 1	
Nom	LECOMTE
Prénoms	Daniel
Rue	157 rue de la Pompe
Code postal et ville	75116 PARIS
Société d'appartenance	MEDIALIVE
Inventeur 2	
Nom	CAPOROSI
Prénoms	Jérôme
Rue	23/25 avenue du Général Leclerc
Code postal et ville	92340 BOURG-LA-REINE
Société d'appartenance	
DATE ET SIGNATURE(S) DU (DES) DEMANDEUR(S) OU DU MANDATAIRE	

La loi n°78-17 du 6 janvier 1978 relative à l'informatique aux fichiers et aux libertés s'applique aux réponses faites à ce formulaire. Elle garantit un droit d'accès et de rectification pour les données vous concernant auprès de l'INPI.

**PCT/FR2004/000947**

